DOCUMENT RESUME

ED 128 678 CG 010 782

AUTHOR Riggs, Donald E.
TITLE Improving Reliability and Validity of "Achievement via Conformance" Through Computer Applications.
PUB DATE [71]
NOTE 11p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Academic Achievement; Conformity; Evaluation Methods; Grade Point Average; Higher Education; *Item Analysis; Psychological Studies; *Psychological Testing; *Test Reliability; *Test Validity
IDENTIFIERS *Achievement via Conformance Scale; *California Psychological Inventory

ABSTRACT

This paper describes an experiment conducted in order to improve the reliability and validity of the Achievement via Conformance (AC) scale of the California Psychological Inventory (CPI). The primary goal of AC is to identify those factors of interest and motivation which facilitate achievement in any setting where conformance is positive behavior. The purposes of the study were: (1) to examine the correlation of the AC scale with grade point average for 40 randomly selected college male freshmen; and (2) to attempt to improve the reliability and validity of the AC scale by dropping various test items after each computer analysis. It was found that elimination of "bad" test items of the scale enhanced both its validity and reliability. Relationships between AC and various intellectual variables were found to be in need of much additional empirical investigation. (SJL)

IMPROVING RELIABILITY AND VALIDITY OF "ACHIEVEMENT VIA CONFORMANCE" THROUGH COMPUTER APPLICATIONS

Dr. Donald E. Riggs
Director of Libraries and Media Services
Bluefield State College, Concord College and Greenbrier Community College

Introduction. This writer has always been intrigued with and interested in the various psychological tests and their relationships with measures of intelligence. Having had some experience with psychological tests (e.g., Ohio State Psychological Examination) and an enthusiasm to learn more about various aspects of psychological testing, this project provided an opportunity to experiment with the improvement of reliability and validity of the Achievement via Conformance (Ac) scale of the California Psychological Inventory (CPI).

The primary goal of Ac is to identify those factors of interest and motivation which facilitate achievement in any setting where conformance is a positive behavior. High scorers tend to be seen as: capable, co-operative, efficient, organized, responsible, stable, and sincere; as being persistent and industrious; and as valuing intellectual activity and intellectual achievement. Low scorers tend to be seen as: coarse, stubborn, aloof, awkward, insecure, and opinionated; as easily disorganized under stress or pressures to conform; and as pessimistic about their occupational futures.

It is interesting to note the apparent contradictions in the results of various studies concerning the usefulness of Gough's (4) California Psychological Inventory (CPI) as a predictor of academic achievement. In particular, the Ac (Achievement via Conformance), Ai (Achievement via Independence), and Ie (Intellectual Efficiency) scales have been investigated in relation to grade point average (GPA) and several measures of intellectual ability. The outcomes of such studies have been generally inconsistent and equivocal. For example, Bendig (1)

obtained correlations running from .35 to .68 for men and from .42 to .46 for women in his study of Ac as a predictor for GPA in psychology courses. Watson (9) also experimented with the three intellectual related scales of the CPI.

Purpose of the Study. The purposes of the study were: 1.) To examine the correlation of the Ac scale of the CPI with GPA for forty randomly selected college male freshmen who had completed at least the equivalent of one full-time semester (12 credit hours) at Bluefield State College; 2.) To attempt to improve the reliability and validity of the Ac scale by dropping various test items after each computer run; 3.) To obtain some personal satisfaction and enjoyment from working with a psychological testing instrument which would result in a rewarding learning experience.

Procedure. Forty male freshmen were selected for the experiment with the Ac scale (32 test items) by taking the first forty male students in order of social security numbers (i.e., in an attempt to get a random sample). The GPA's of these students were used as the criterion scores and the correct responses of the Ac scale were used as the predictor scores. These scores were then programmed via BMD02D Correlation with Transgeneration Package. The package enabled one to obtain a computer printout of sums, means, standard deviations, correlation matrices, and a plotting graph of variables. Computations were performed utilizing: 1.) The KR 20 Formula $\frac{n}{n-1}\left(\frac{1-\sum_{i=1}^{32}\sigma_i^2}{\left(\sum_{i=1}^{32}\sigma_i r_{ix}\right)^2}\right)$; 2.) Correlation coefficients for validity; 3.) A validity graph (see Table II in Appendix); 4.) A reliability graph (see Table III in Appendix); 5.) A difficulty factor for each item (see Table IV in Appendix).

Findings.

1. Table I (Appendix) displays the computer printout and computations of the first computer run. Results of the first computer run: KR 20 = .7413, r_{xy} = .6504.

2. Table II (Appendix) exhibits the plotting of the validity with the centroid ($\sigma_x r_{iy} \div 32$ and $\sigma_i r_{ix} \div 32$). The results on the graph reveal that items number 14, 16, 17, 19, 20, 22, 25, 27 and 30 should be dropped in order to improve validity in the second computer run. Results of the second computer run: KR 20 = .7687, r_{xy} = .6612.

3. Table III (Appendix) exhibits the plotting of the reliability. The results on the graph reveal that items number 8, 10, 12, 13, 15, 17, 19, 20 and 27 should be dropped in order to improve reliability in the third computer run. Results of the third computer run: KR 20 = .7679, r_{xy} = .6748.

4. A fourth computer run was made with the intent in mind to eliminate identical items (17, 19, 20, 27) which would simultaneously enhance both validity and reliability as compared with the first computer run. Results of the fourth computer run: KR 20 = .7510, r_{xy} = .6646.

Discussion. Computer run two is best for improving validity; computer run three gives the best results for reliability. As denoted in the findings, the Ac scale's reliability and validity can both be improved simultaneously (computer run four). This particular method of improving both of these aspects at the same time is of signal importance in test construction and refinement.

Relative to the difficulty factor of each test item (Table IV - Appendix), one should note that the extreme differences in difficulty of some items would most likely hinder one from applying the KR 21 Formula to the results of this particular Ac scale. Essentially, the KR 21 Formula would most likely yield lower approximations if applied to this data than would KR 20.

For purely curiosity's sake, a criterion score (GPA) and a predictor score (correct responses of a 50 item test) of the Intellectual Efficiency (Ie) scale of the CPI were obtained in a separate computer run. The KR 20 was .7189 and r_{xy} was .6200. Further research is needed to be done with two-scale analyses (e.g.,

Ac and Ie) of the CPI. Finding the intercorrelation of the scales would provide an interesting challenge.

Conclusion and Recommendations. One thing is evident from this study: By eliminating "bad" test items of the Ac scale both validity and reliability may be enhanced. Gough (5) has revealed that the Ac scale has positive correlation with academic achievement. However, based on a close scrutiny of the literature, this writer could not find any trace of where anyone has attempted to refine the CPI by eradicating some of the "bad" items of the various scales. Undoubtedly, some of the same items will continue to appear to "fall short" during the various administrations of the CPI. Why has someone not restructured the CPI? If the CPI in its entirety, or in parts by scales, is to contain the utmost reliability and validity relative to test items, then constant refinement needs to be done.

From this project and investigation of the literature, it can only be concluded that the relationships between Ac and various intellectual variables require much additional empirical investigation before a more concrete understanding may prevail. Perhaps enough interest has been stimulated by this project to encourage someone to independently explore more thorough means of improving the efficiency of the Ac and other scales as positive indicators.

SELECTED REFERENCES

1. Bendig, A. W. "Comparative Validity of Empirical Temperament Test Keys in Predicting Success in Psychology," Journal of Educational Research, LI (1958), pp. 341-348.

2. Cronbach, L. J. "Test Reliability: Its Meaning and Determination," Psychometrika, XII (1947), pp. 1-15.

3. Engel, B. M. "Academic Achievement and Nongradedness," Journal of Experimental Education, IL (1971), pp. 24-26.

4. Gough, H. G. The California Psychological Inventory (Palo Alto, California: Consulting Psychologists Press, 1960).

5. Gough, H. G. "Academic Achievement in High School as Predicted from the California Psychological Inventory," Journal of Educational Psychology, LV (1964), pp. 174-181.

6. Holland, J. L. "The Prediction of College Grades from the California Psychological Inventory and the Scholastic Aptitude Test," Journal of Educational Psychology, L (1959), pp. 135-142.

7. Kuder, G. F. and Richardson, M. W. "The Theory of the Estimation of Test Reliability," Psychometrika, II (1937), pp. 151-160.

8. Runyon, R. P. Fundamentals of Behavioral Statistics (Reading, Mass.: Addison-Wesley, 1967).

9. Watson, C. G. "The California Psychological Inventory as A Predictor of Academic Achievement in Normal and Maladjusted College Males," Journal of Educational Research, I (1967), pp. 10-13.

APPENDIX

TABLE I. COMPUTER PRINTOUT AND COMPUTATIONS

Items	Variance σ_i^2	Standard Deviations σ_i	Correlation With Predictor r_{ix}	$\sigma_i r_{ix}$	Correlation With Criterion r_{iy}	$\sigma_i r_{iy}$
1	.2045	.4522	.3999	.1808	.3232	.1461
2	.1122	.3349	.2532	.0848	.2971	.0995
3	.2506	.5006	.5159	.2583	.3450	.1727
4	.2461	.4961	.4777	.2370	.3907	.1938
5	.1641	.4051	.3869	.1567	.3216	.1303
6	.2564	.5064	.5379	.2724	.3744	.1896
7	.1923	.4385	.4522	.1983	.4174	.1830
8	.2461	.4961	-.1984	-.0984	-.2163	-.1073
9	.2154	.4641	.5127	.2380	.3394	.1575
10	.2154	.4641	-.0432	-.0200	-.2083	-.0967
11	.1481	.3848	.4278	.1646	.6146	.2365
12	.1923	.4385	.1580	.0693	.1358	.0595
13	.1481	.3848	.1186	.0456	.0767	.0295
14	.2557	.5057	.5796	.2931	.2485	.1257
15	.2045	.4522	-.0724	-.0327	-.1715	-.0775
16	.2558	.5057	.5985	.3027	.2613	.1321
17	.2154	.4641	.1730	.0803	.0750	.0348
18	.1788	.4229	.4197	.1775	.5452	.2306
19	.1923	.4385	.2016	.0884	.0919	.0403
20	.1481	.3848	.2663	.1025	-.0153	-.0059
21	.2538	.5083	.5434	.2738	.2935	.1479
22	.2557	.5057	.5163	.2611	.2067	.1045
23	.1641	.4051	.5803	.2351	.2918	.1182
24	.0711	.2667	.3412	.0909	.3396	.0906
25	.1481	.3848	.7767	.2989	.3015	.1160
26	.1307	.3616	.7650	.2766	.4633	.1675
27	.2045	.4522	.4454	.2014	.0957	.0433
28	.1122	.3349	.3388	.1135	.3882	.1300
29	.2045	.4522	.3260	.1474	.4833	.2185
30	.1923	.4385	.3759	.1648	.1397	.0613
31	.2461	.4961	.5721	.2838	.4804	.2383
32	.2564	.5064	.4341	.2198	.3564	.1805
Totals	6.2817	14.0481	12.3791	5.3663	8.0865	3.4907

TABLE II. - VALIDITY

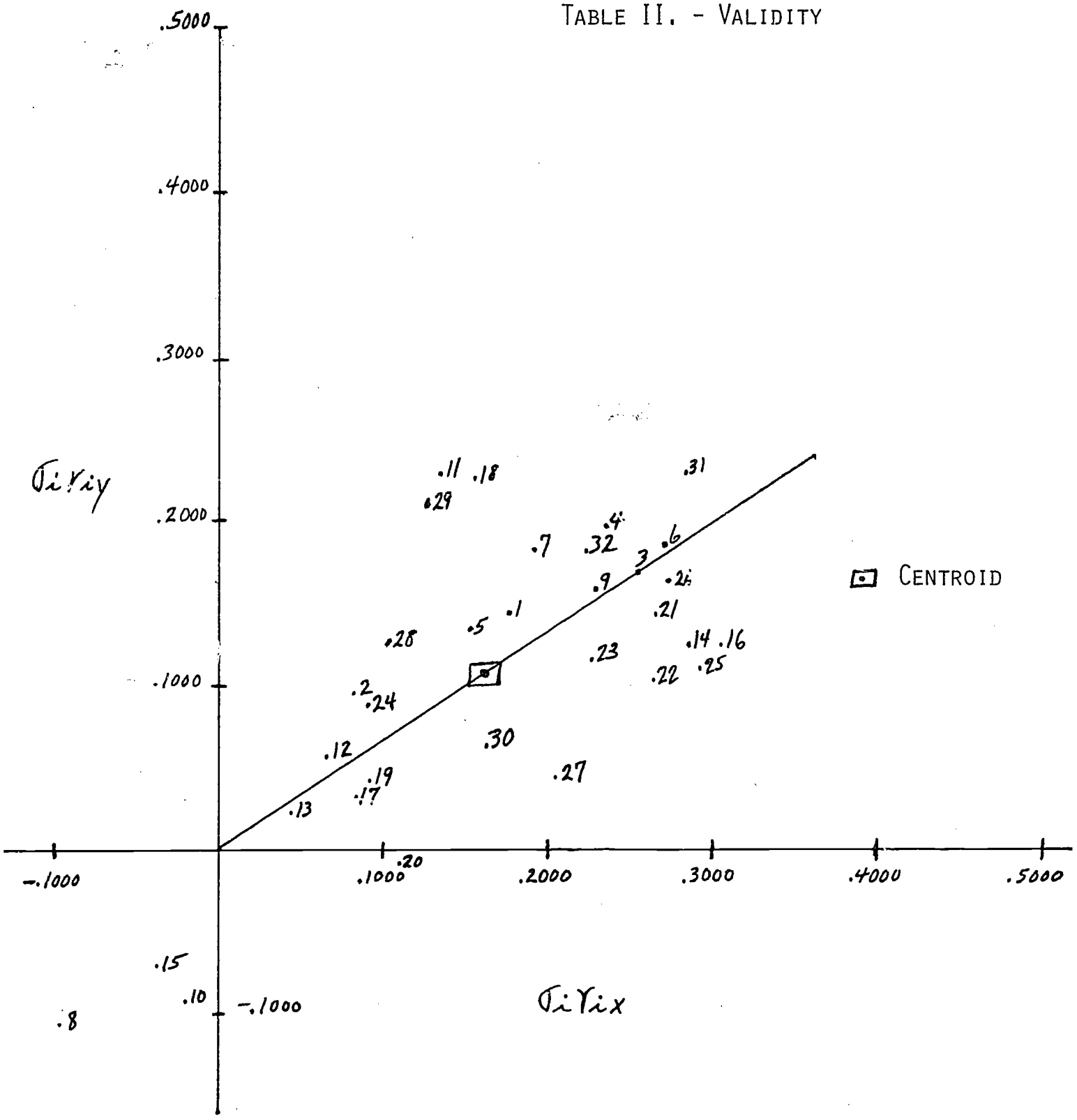

TABLE III. - RELIABILITY

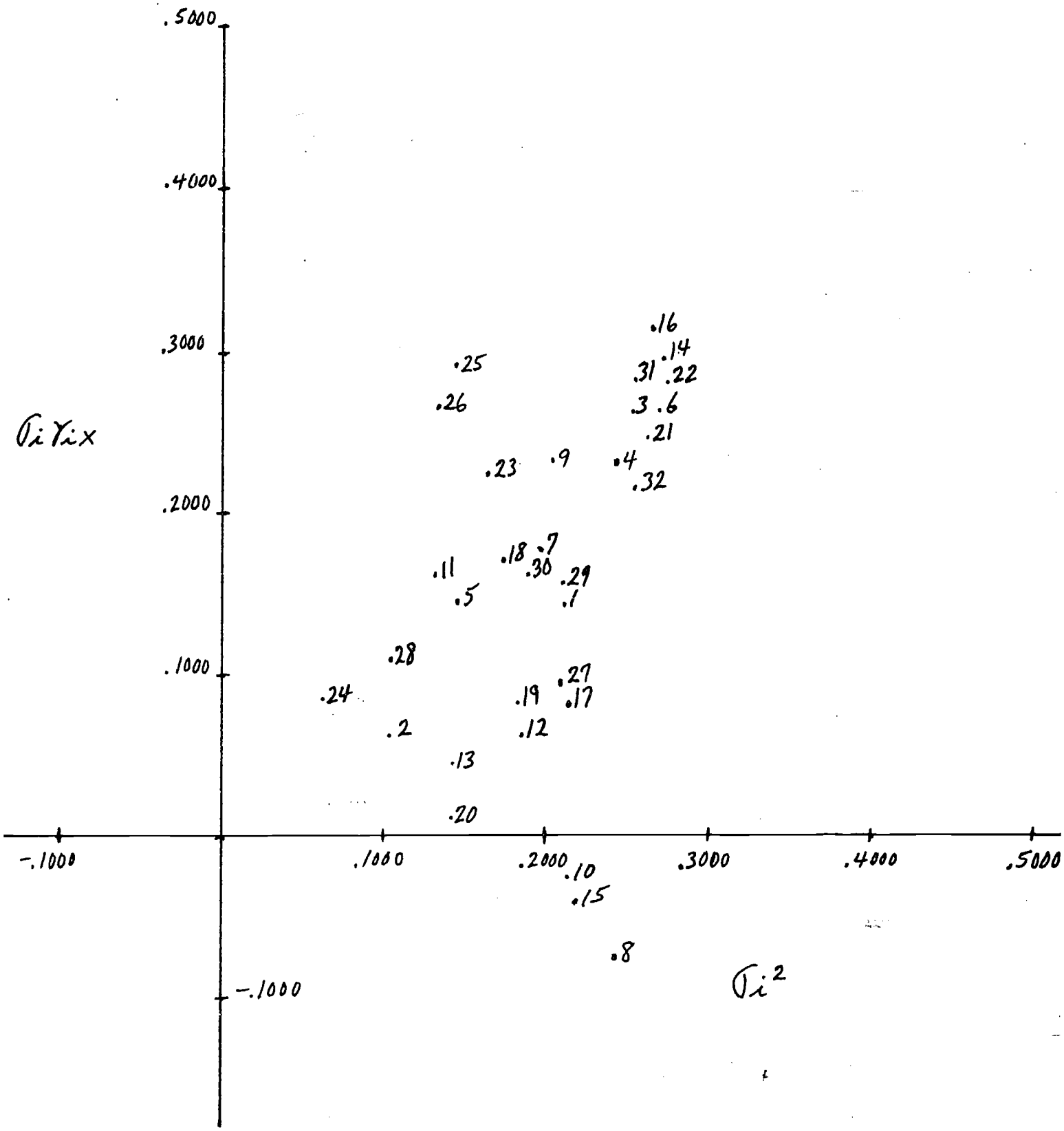

TABLE IV

Items, Number of Correct Responses, and Difficulty Factor in KR20 Calculation

Item	Number of Correct Responses	Difficulty Factor	Item	Number of Correct Responses	Difficulty Factor
1	29	.7250	17	12	.3000
2	35	.8750	18	31	.7750
3	17	.4250	19	10	.2500
4	24	.6000	20	33	.8250
5	32	.8000	21	22	.5500
6	20	.5000	22	21	.5250
7	30	.7500	23	8	.2000
8	16	.4000	24	37	.9250
9	12	.3000	25	7	.1750
10	12	.3000	26	6	.1500
11	33	.8250	27	11	.2750
12	30	.7500	28	35	.8750
13	7	.1750	29	29	.7250
14	19	.4750	30	10	.2500
15	11	.2750	31	16	.4000
16	19	.4750	32	20	.5000